SHIVAM LALAKIYA

Boston, MA | 781-363-8482 | lalakiya.s@northeastern.edu | LinkedIn | portfolio

EDUCATION

Northeastern University, College of Engineering | Boston, MA   **May 2023 (Expected)**
Master of Science in Data Analytics Engineering   GPA: 3.81/4.0
<u>Relevant Courses</u>: Foundation of Data Analytics, Data Management, Computation and Visualization, Machine Learning, NLP

Sardar Vallabhbhai National Institute of Technology | Surat, India   **May 2021**
Bachelor of Technology in Electronics and Communication Engineering   CGPA: 8.10/10.0

TECHNICAL SKILLS

Languages/ Frameworks:	Python, R, SQL, C/C++, MongoDB, Pytorch, Tensorflow, Keras, Shell scripting, Perl, MATLAB, NoSQL
Tools and Technologies:	NLP, Neural Networks, Deep Learning, Computer Vision, RNNs, Power BI, Tableau, ETL, Transformer, Encoder-Decoder, AWS, git, CI/CD, Snowflake, Kafka, Spark, Hadoop, RDBMS, Hive, Hypothesis Testing
Libraries:	Scikit-Learn, Pandas, Numpy, OpenCV, TensorFlow, Matplotlib, Seaborn, NLTK

EXPERIENCE

Ring Therapeutics | Data Science co-op | Boston, MA   **May 2022 - Present**
- Provided analytical insights regarding binding sites, tissue, and cell specificity from sequenced data collected by the Discovery team from the patients' protein sequences to enhance the virology team's outcome and develop Anello-backed programmable medicine.
- Leveraged NLP-based models for genes/protein sequences to predict tropism to create viral vectors that can safely and effectively deliver therapies to target cells and tissues.
- Created ETL pipeline to load experiment data into fasta format using Airflow DAG to eliminate manual data loading to the server.

Northeastern University | Course Assistant, Introduction to Distribution Intelligence   **Jan 2022 – May 2022**
- Collaborated with Prof. Milad Siami to design the assignment and homework for Introduction to Distributed Intelligence.
- Guided 50+ students with courses and assignments and conducted five office hours to ensure learner success and course completion.

IIT Madras | Research Assistant | Chennai, India   **April 2020 - Sept 2020**
- Built a project, 'Caching using Deep Learning,' which involved time-series prediction based on trends using RNNs.
- Trained a custom dataset of size ~12GB filtered using Pandas.
- Expanded this LSTM-based caching policy, where RNNs were used to predict user requests' future preferences with 90% accuracy.
- Achieved 130% better hit rates than existing caching policies such as LIFO, LRU, and LFU.

Sardar Vallabhbhai National Institute of Technology | Research Intern | Surat, India   **Jul 2019 – Dec 2019**
- Administered a project entitled "Multimodal Biometric System,", where Iris, Facial, Speech Recognition, and fingerprints were combined, utilized, and implemented using CNNs to obtain 85% better precisions create a reliable biometric system.
- Applied this model in the department for attendance system, which saved about 15 minutes for taking attendance in each class every day and implemented it in the professor's cabin to enhance security.

ACADEMIC PROJECTS

Stock Market sentiment and time series analysis (Python, LSTM, Streamlit, Web-scrapping, Java)   **May 2022 – Aug 2022**
- Extracted recent 30 days' stock data from WSJ and historical prices from Tiingo API using ticker symbols to store them in CSV format.
- Developed the Naïve-Bayes model to predict trends and future prices using the LSTM model with more than 90% accuracy.
- Created interactive front-end utilizing Streamlit library to find insights into any listed stock and reach a mass audience.

Flight Booking Website (Javascript, CSS, HTML, NodeJs, CI/CD, REST API, Docker)   **Sep 2021 – Dec 2021**
- Created a flight booking website that fetches live flight ticket information from sky scanner API using React.JS, NodeJs for backend infrastructure containing controllers, and MongoDB to persist data.
- Built User, Admin, and Airline modules with REST APIs for performing different CRUD operations like creating and updating user, flight, airline information, as well as creating and applying deals to different flights for different users.
- Deployed the React framework in Netlify and the NodeJs portion in Heroku and also dockerized the application to run in a remote environment and support CI/CD pipeline.

Analytics and Visualization using R programming (R, SQL, Tableau, ETL, Time-series forcasting)   **Sep 2021 – Dec 2021**
- Performed clustering, probabilistic analysis, and text mining to find insights on University and E-commerce datasets.
- Derived the most affecting factors for ranking the top 100 universities and reasons behind the changes in hierarchy.
- Obtained the correlation between discounts, sales, and profit for e-commerce websites and concluded that festive seasons have 35% higher sales and non-festive seasons have 30% higher deals.

Hyperspectral Image Classification using Deep Learning (Python, Keras, OpenCV, Hypothesis testing)   **Jul 2020 – May 2021**
- Collaborated with Dr. Jigish Patel to implement Hyperspectral Image Classification using CNNs, GANs, and PCA.
- Worked on PCA and k-PCA to reduce the time and space complexity by 50%, projecting higher dimensions using kernel methods.
- Compared the existing CNN algorithm with PCA-CNN and obtained at least 30% less training time and complexity.